Statistical Analysis Plan (SAP)

Protocol Number: NEU_2020_02

Single-Arm, Prospective, Multicenter Study to Evaluate the Effectiveness and Safety of the NeuWave Certus Microwave Ablation System in Chinese Patients with Primary or Secondary Tumors in the Lung

Protocol Version: v2.0, 15 September2021

SAP Version 1.0

SAP Version Date: March 22, 2022

The following individuals have reviewed this version of the Statistical Analysis Plan and are in agreement with the content:

Signature Page

DMed Biostatistician Approver:

Jack Li	Jack. Li	01 Apr 2022
(Print)	(Sign)	Date

Sponsor Study Biostatistician:

Yun Kathy Chen	Refer to vTMF	Refer to vTMF
(Print)	(Sign)	Date

Head of Biostatistics:

Srichand Jasti	Refer to vTMF	Refer to vTMF
(Print)	(Sign)	Date

Franchise Clinical Study Lead:

Erin Meyers	Refer to vTMF	Refer to vTMF
(Print)	(Sign)	Date

Franchise Clinical Platform Lead :

Patricia Schleckser	Refer to vTMF	Refer to vTMF
(Print)	(Sign)	Date

Revision History

Revision Number	Revision Date (DD/MMM/YYYY)	Reasons for Revision
Version 1.0	22/Mar/2022	NA

Abbreviations

AE	Adverse Event
CI	Confidence Interval
CP	Conditional Power
CRF/e-CRF	Case Report Form / Electronic Case Report Form
CSR	Clinical Study Report
CT	Computed Tomography
DMC	Data Monitoring Committee
ECOG PS	Eastern Cooperative Oncology Group Performance Status
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire -Core 30
EORTC QLQ-LC13	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire -Lung Cancer-specific Module
FAS	Full Analysis Set
GGO	Ground Glass Opacity
ICF	Informed Consent Form
IRC	Independent Review Committee
KM	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
NMPA	National Medical Products Administration
MW	Microwave
NPRS	Numeric Pain Rating Scale
NSCLC	Non-small Cell Lung Cancer
OS	Overall Survival
PD	Progression of Disease
PFS	Progression-Free Survival
PG	Performance Goal
PI	Principle Investigator
PP	Per Protocol
PT	Preferred Term
QOL	Quality of Life
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SE	Standard Error
SOC	System Organ Class

Table of Contents

This Statistical Analysis Plan (SAP) describes the analyses and data presentations for protocol NEU_2020_02 (protocol version 2.0, 15September2021).

This document will serve as the final guidance for all the statistical analyses for this study and will supersede the Statistical Considerations section in the protocol if there are any discrepancies. Any deviation from this SAP will be documented in the clinical study report (CSR).

1 Study Overview

1.1 Study Design

This study is a single-arm, prospective, multicenter clinical trial with a performance goal (PG).

The study population includes adult subjects with non-small cell lung cancer (NSCLC) or oligometastatic lung tumors who plan to receive percutaneous microwave ablation. Individuals who have lung tumor(s) meeting the eligibility criteria of the study and sign the Informed Consent Form will be enrolled into this study.

Each enrolled subject will have his or her eligible lung tumor(s) ablated using the NeuWave Certus Microwave Ablation System (in conjunction with the NeuWave Certus Microwave Ablation Probes) in a single procedure. Subjects will be followed at 1, 3, 6, and 12 months after the initial microwave ablation procedure to evaluate the effectiveness and safety of percutaneous microwave ablation with the study device.

To provide study sites with an opportunity to get equal experience in the use of the Certus system, each study site will have two (2) subjects treated as part of a run-in phase. These subjects will only be included in the safety analysis set.

1.2 Study Objective

The objective of this study is to evaluate the effectiveness and safety of the NeuWave Certus Microwave Ablation System and Accessories for percutaneous microwave ablation of non-small cell lung cancer and oligometastatic lung tumors.

1.3 Study Endpoints

1.3.1 Primary Endpoint

Technical Efficacy Rate: Percentage of tumors that are completely covered by the ablation area without signs of pathological enhancement as assessed by lung enhanced CT at Visit 3 (i.e., 30 days±7 days after the first ablation).

1.3.2 Secondary Endpoints

- Secondary Efficacy Endpoints

- Technical Success Rate: Percentage of tumors that achieve A0 or A1 ablation classification determinations (i.e., complete tumor ablation with a surrounding margin) based on the lung CT immediately following the initial ablation procedure.
- A0 Ablation Rate: Percentage of tumors that achieve A0 ablation (i.e., the ablation zone covers the tumor completely and has a minimal margin of at least 5mm for secondary lung tumors and at least 10mm for primary lung tumors) based on the lung CT immediately following the initial ablation procedure.

  Note: NSCLC are primary tumors and oligometastatic lung tumors are secondary.
- Re-Ablation Technical Success Rate: Percentage of tumors that achieve Technical Success in the lung CT immediately following the repeat ablation with the study device during the study.
- Local Tumor Progression: Local tumor progression and time to local tumor progression of any original-ablated tumor(s).
- Progression-Free Survival: Length of the time the subject is still alive after the original ablation procedure and with no evidence of any tumor progression (local, regional, or distant).
- Overall Survival: Length of time that the subject is still alive after the original ablation procedure within the study duration.

- Safety Endpoints
  - All AEs from the start of any ablation procedure (starting from the time of probe puncture on skin) to 30 days post-ablation or early discontinuation.
  - All SAEs from the start of ablation procedure through the end of the study or early discontinuation.
  - All device-related AEs, procedure-related AEs, device-related SAEs, and procedure-related SAEs from the start of ablation procedure through the end of the study or early discontinuation.

1.3.3 Exploratory Endpoints

- Patient-Reported Outcomes
  - Ablation-Procedure Related Pain: Compare the 1 day, 3 days, and 1-month post-ablation versus pre-ablation patient-reported pain levels using the Numeric Pain Rating Scale.
  - Quality of Life: Compare the patient's 1 month post-ablation, 6 months post-ablation, 12 months post-ablation to the pre-ablation Quality of Life scores using EORTC QLQ-C30 and Lung-specific QLQ-LC13.

- ECOG Performance Status: Patient functionality as measured by distribution of Eastern Cooperative Oncology Group (ECOG) classification scores over time.
- Ablation Procedure Related Indicators: Number and types of probes used, number of ablation cycles, ablation power, and ablation time.
- Length of Hospital Stay: The duration from the patient's completion of ablation procedure to discharge.
- Re-admission Rate: Percentage of unplanned hospital admission or re-admission due to an adverse event occurring within 30 days after any ablation procedure using the study device (i.e., the original ablation or a repeat ablation).

2 Treatment Assignment

All enrolled subjects will undergo percutaneous microwave ablation with the NeuWave Certus Microwave Ablation System (in conjunction with NeuWave Certus Microwave Ablation Probes).

3 Randomization and Blinding Procedures

This is a single-arm, open-label study. No randomization will occur.

The study team will be blinded to the summary results of the primary effectiveness endpoint until the database lock of the study primary analysis. Details of blinding procedures are documented in the blinded plan.

4 Interval Windows

Interval windows are provided in the Schedule of Activities in Table 1 of the protocol. Data collected in Unscheduled Visit forms will be listed as such.

5 Levels of Significance

In this study, the statistical power of 80% and the significance level of one-sided 0.025 are used for the sample size calculation based on the primary effectiveness endpoint (i.e. Technical Efficacy Rate) hypothesis testing.

6 Analysis Sets

The analysis sets are defined as follows:

- The Full Analysis Set (FAS) is defined as all subjects who are enrolled in the study and received ablation power (excluding the run-in subjects).
- The Per Protocol (PP) set is defined as subjects in FAS without major protocol deviations that have great impact on the primary effectiveness endpoint.

- The Safety Analysis Set (SAF) includes all subjects who provide informed consent and have microwave ablation attempted with the NeuWave system (including the run-in subjects).
- Run-in Set includes all subjects who are enrolled in the study and enter the run-in phase.

The primary effectiveness endpoint will be analyzed on the FAS and PP. The analysis on the FAS is the primary analysis and the analysis on the PP set is the sensitivity analysis.

The secondary and exploratory effectiveness endpoints will be analyzed on the FAS.

All safety endpoints will be analyzed on the SAF. The AE summary table will be presented on the FAS and run-in set as well.

For subjects included in the run-in set, all baseline data, surgical information, and effectiveness data will be described as the continuous variable and categorical variable. No confidence interval (CI) will be provided. Only observed data will be analyzed and no missing data imputation rule will be applied to subjects in the run-in set.

The protocol deviations impacting the PP set will be reviewed and finalized during the data review meeting held prior to database lock of the primary analysis.

7 Sample Size Calculation

This study plans to enroll 120 subjects plus run-in subjects (each site will enroll 2 run-in subjects). The sample size calculation is justified as follows:

The study sample size is determined by hypothesis testing for the comparison of the primary endpoint versus the performance goal (PG). The study's PG is set as 80%. The primary hypothesis test of the primary endpoint is: H0: P≤ P0 (80%) vs. H1: P>80%, where P is the actual Technical Efficacy Rate of the Neuwave Certus Microwave Ablation System. A sample size of 120 subjects is necessary to achieve at least 80% power for demonstrating that the Technical Efficacy rate of the NeuWave Certus Microwave Ablation System is > 80% when the expected performance is 90% and a one-sided significance level of 0.025 is used with the Normal Approximation to the Binomial distribution (Z-test that uses S(P0) to estimate the standard deviation), with adjustment for up to 10% dropout.

Each site will enroll 2 subjects first in the run-in phase. So the total sample size will be 120 subjects plus run-in subjects (each site will enroll 2 run-in subjects) .

Of the 120 non run-in sample, the number of cases for either disease type is limited at up to 75% (i.e., in the 120 non run-in sample, up to 90 subjects could be NSCLC or oligometastatic lung tumor).

In addition, given the lack of available clinical data for the evaluation of the performance of the investigational product in pulmonary tumor ablation procedures, an adaptive approach will be used to re-estimate the sample size as per the "promising zone" method of Mehta and Pocock (2011) after 50% of the target enrolled subjects complete the Visit 3 Technical Efficacy Evaluation. This procedure involves evaluation of the conditional power (CP) in the interim analysis. If the CP is to fall in the pre-specified "promising zone", the sample size will

be increased to boost the CP up to the target level of 80% for the primary effectiveness endpoint, in the meanwhile, the new sample size will be subject to a predefined maximum value (200 enrolled subjects). If the CP is to fall out of the pre-specified “promising zone”, the study will maintain the originally planned sample size, i.e. 120 subjects plus run-in subjects. For details in sample size re-estimation method, please refer to Section 8.12.

8 Statistical Analyses

8.1 General Conventions

In addition to listing variables of interest, analysis methods for this study will focus on summaries using descriptive statistics appropriate to the type of variable under consideration.

- Categorical variables will be presented as frequencies and percentages.
- Continuous variables will be presented with the number of values, the number of missing, mean, standard deviation, median (Q1, Q3), minimum, and maximum values.
- For percentage calculation, the number of missing data will be not included in the denominator unless otherwise specified.

All listings, including scheduled visits and unscheduled visits, will be listed on all SAF subjects, and sorted by stage (run-in, non run-in phase), site ID, and subject ID unless otherwise specified.

All analyses will be conducted using SAS version 9.4 or higher.

8.2 Disposition of Study Subjects

Subject disposition will be presented for all screened subjects.

The number of subjects screened (failed, screen failure reasons, enrolled), in each analysis set, those who received ablation and those who complete and discontinue the study along with the specific reasons for discontinuation will be tabulated in total and also by site.

Identified lesion to be ablated disposition will be summarized by target lesion (yes, no) using counts in the SAF.

Listings will be provided and sorted by screening status (failed, enrolled), screening failed reason, phase (run-in, non run-in phase), site ID, subject ID, date of ablation, subject status, and date of study completion/early discontinuation.

8.3 Protocol Deviation

Protocol deviations will be presented for all enrolled subjects.

All protocol deviations and major protocol deviations will be tabulated by categories. Number of subjects and percentage for each category will be provided.

All protocol deviations will be listed and sorted by type (major, minor), site ID, subject ID, and deviation sequence number.

8.4 Demographics and Baseline Characteristics

Analyses will be conducted on the SAF, FAS, and run-in set.

Demographics and other baseline characteristics will be summarized using descriptive statistics. Continuous variables include age and vital signs, including weight, height, and BMI [derived as weight (kg) divided height^2 (m^2)]. Categorical variables include age category (≤50 years, >50 years), gender (male, female, for females: of childbearing potential, permanently sterilized, postmenopausal), ethnicity group, race, ECOG-PS score, smoking status, number of target lesions (identified and ablated) (1,2,3), number of identified lesion to be ablated but not ablated (0,1,2,3), number of identified lesion not to be ablated (0,1,2,3,4,5), subject target tumor diagnosis (stage IA1 NSCLC, Stage IA2 NSCLC, Oligometastatic lung tumor, other).

Listings will be provided.

8.5 Medical History/ Surgical history

Analyses will be conducted on the SAF, FAS, run-in set

Medical history/Surgical history will be summarized by each category and also by presenting number and percentage of participants by System Organ Class (SOC) and Preferred Term (PT) using MedDRA; each participant could have medical history under multiple SOC and PT, but each participant will be counted only once within each category, each SOC and PT.

Listings will be provided.

8.6 Radiotherapy History

Analyses will be conducted on the SAF, FAS, run-in set.

Radiotherapy history will be summarized by treated lesion types (primary lung tumor related, secondary lung tumor patient‘s original lesion related, secondary lung tumor patient‘s metastatic lesion related, and other).

Listings will be provided.

8.7 Prior and Concomitant Medications/Procedures

Analyses will be conducted on the SAF, FAS and run-in set.

Prior medications/procedures will be defined as medications/procedures that start prior to the date of initial ablation. Concomitant medications/procedures will be defined as medications/procedures that start on or after the date of initial ablation, or start before initial ablation but still ongoing use on the date of initial ablation.

Prior/concomitant medication/procedures will be summarized by medication category/procedure type and indication that medication/procedure taken for, respectively.

The medication categories include:

- Blood-thinning/anticoagulants
- Anti-neoplastic agents
- Steroids/Anti-inflammatory drugs/ NSAIDs
- Antibiotics
- Others

In addition, the indication categories include:

- Prophylactic
- Adverse Event
- Medical History
- Disease under investigation
- Other

Listings will be provided.

8.8 Lesion assessment

Analyses will be conducted at the target tumor level on the SAF, FAS, and run-in set.

For target tumors, defined as the tumors that have been identified to be ablated and receive ablation power during the initial ablation procedure, the following tumor characteristics will be summarized:

- lung tumor type (primary lung tumor, secondary lung tumor, other), type of ground glass opacity (GGO) (pure GGO, partial GGO, pure solid tumor).
- tumor diagnosis category, location of original tumor, and any other metastatic site (for oligometastatic lung tumor only).
- number, location, and size (the size measured nearest to the ablation date will be used if size is collected more than once) of the tumor to be ablated.
- For primary and secondary tumors, biopsy results (pathology type) of the lesion to be ablated.

All identified lesions to be ablated at visit 1 will be listed on the SAF and sorted by phase (run-in, non run-in), site ID, subject ID, target tumor indicator (target tumor [i.e. tumor identified and ablated], tumor identified but not ablated), and tumor ID, ablation date, tumor type.

For identified lesions not to be ablated, summary statistics will be provided for number, location, and size of tumors in the lung on all enrolled patients.

8.9 Primary Endpoint and Associated Hypotheses

Analyses will be conducted on the FAS, PPS, and run-in set.

The primary efficacy endpoint is Technical Efficacy Rate, defined as the percentage of tumors that are completely covered by the ablation area without signs of pathological enhancement as assessed by lung enhanced CT at Visit 3 (i.e., 30 days±7 days after the first ablation). Each subject will have a maximum of one NSCLC tumor or a maximum of three ipsilateral oligometastatic lung tumors to be ablated, each of which will be defined as a success or failure. If the initial ablated tumor had re-ablation before Visit 3 Technical Efficacy assessment, the re-ablated tumor will be considered as a failure. The re-ablation will not change or extend the subject's follow-up schedule. Technical Efficacy Rate will be analyzed based on both the target tumor level and subject level.

The Technical Efficacy Rate is assessed by both the study physician (i.e., PI) and Independent Central Imaging Committee (IRC). Primary statistical analysis and hypothesis testing for the primary effectiveness endpoint will be based on IRC's evaluation results. For IRC evaluation, there will be at least two reviewers providing their results. If the two reviewer's results are the same, the first one's result will be used. If the two reviewer's results differ, a third reviewer will evaluate and his or her result will be used for the analysis. This analysis rule will be applied for all IRC-based endpoint.

The null and alternative hypotheses for primary analysis of the primary endpoint are as follows:

H_0: P ≤ P0 (80%)

H_1: P > 80%,

where P is the actual Technical Efficacy Rate using the NeuWave Certus Microwave Ablation System.

If the initial ablated tumor had re-ablation before the Visit 3 Technical Efficacy assessment, the re-ablated tumor will be considered as a failure. Other tumors with missing parameters or other early discontinuations will not have data imputation.

For multi-target-tumor subjects, his or her largest Target Tumor, determined by the maximum three dimensions of each target tumor, will be used for subject-level analysis conservatively.

The number and proportion of tumors/subjects achieving Technical Efficacy at Visit 3 will be summarized and the corresponding 95% CI based on the Normal Approximation to Binomial Distribution (Z-test that uses S(P0) to estimate the standard deviation) or Clopper-Pearson method (if the proportion is greater than 90%) will be calculated.

The success criterion of the primary effectiveness endpoint will be met if the lower limit of the two-sided 95% CI for Technical Efficacy Rate is higher than 0.8.

The following additional sensitivity analyses will be conducted for the primary effectiveness endpoint using same statistical method:

- Sensitivity analysis 1: based on the PP population at both the target tumor level and subject level.
- Sensitivity analysis 2: based on the FAS and different imputation rule for missing data of the primary effectiveness endpoint (see in Section 8.14) at both the target tumor level and subject level.

Additional analysis: based on the FAS and primary effectiveness endpoint values assessed by the investigator at the target tumor level.

Listings, including IRC assessment and PI assessment, will be provided.

8.10 Secondary Endpoints and associated hypotheses

No hypothesis testing is performed for the secondary endpoints. There is no multiplicity adjusted.

Secondary effectiveness endpoints will be analyzed on the FAS and run-in set unless otherwise specified.

8.10.1 Technical Success Rate

Analyses will be performed at target tumor level.

Technical Success Rate is defined as the percentage of tumors that achieve A0 or A1 ablation (i.e., the ablation zone covers the tumor completely and has a surrounding minimal margin) based on the lung CT immediately following the initial ablation procedure. Technical Success Rate will be assessed by both IRC and investigators.

The number and percentage of tumors achieving Technical Success assessed by the IRC and investigator will be analyzed, respectively, and the corresponding 95% CI based on the Normal Approximation to Binomial Distribution (Z-test that uses S(Phat) to estimate the standard deviation) or Clopper-Pearson method (if the proportion is greater than 90%).

Listings, including IRC assessment and PI assessment, will be provided.

8.10.2 A0 Ablation Rate

Analyses will be performed at target tumor level.

A0 Ablation Rate is defined as the percentage of tumors that achieve A0 ablation (i.e., the ablation zone covers the tumor completely and has a minimal margin of at least 5mm for secondary lung tumors and at least 10mm for primary lung tumors) based on the lung CT immediately following the initial ablation procedure. A0 Ablation Rate assessed by IRC or investigators will be summarized, respectively, and the corresponding 95% CI based on the Normal Approximation to Binomial Distribution (Z-test that uses S(Phat) to estimate the standard deviation) or Clopper-Pearson method (if the proportion is greater than 90%).

8.10.3 Re-ablation Technical Success Rate

Analyses on those with re-ablation will be performed at target tumor level.

Re-ablation Technical Success Rate is defined as the percentage of tumors that achieve Technical Success based on the lung CT immediately following the repeat ablation with the study device during the study. Re-ablation Technical Success will be assessed by both the IRC and investigators. Re-ablation Technical Success Rate assessed by IRC or investigators will be summarized, respectively, and the corresponding 95% CI based on the Normal Approximation to Binomial Distribution (Z-test that uses S(Phat) to estimate the standard deviation) or Clopper-Pearson method (if the proportion is greater than 90%).

Listings, including IRC assessment and PI assessment, will be provided.

8.10.4 Local Tumor Progression

Analyses will be performed at target tumor level.

The number and proportion of target tumors with local tumor progression will be summarized and the corresponding 95% CI based on the Normal Approximation to Binomial Distribution (Z-test that uses S(Phat) to estimate the standard deviation) or Clopper-Pearson method (if the proportion is greater than 90%).

On the run-in set, only the number and percentage target tumors with local tumor progression will be provided.

The following analysis will be performed on the FAS and at target tumor level.

Local tumor progression rates of any original-ablated tumor(s) at 3, 6, 9, and 12 months will be estimated by the Kaplan-Meier method and their corresponding 95% CIs will be calculated using Greenwood's formula for estimating the variance and log-log transformation with back transformation to CIs on the untransformed scale.

Time to local tumor progression duration (determined by the CT imaging date, including unscheduled visits) will be summarized descriptively using the Kaplan-Meier method for median and quartiles with 95% CIs estimated using the Brookmeyer-Crowley method. The

censoring rule is described as below and frequency of tumors (n, %) experiencing local progression events and the censored reasons will be provided.

For original-ablated tumor(s) without LTP during the study follow-up period, the tumors will be censored at the date of last local tumor assessment.

Listings, including target tumor ID, ablation date, IRC assessment and PI assessment, local tumor progression date, duration (weeks) will be provided.

8.10.5 Progression-Free Survival

Analyses will be performed on the FAS and at subject level. On run-in set, only the number and percentage of subjects with progression or death will be provided.

Progression-Free Survival (PFS) is defined as the time that the subject is still alive after the original ablation procedure and with no evidence of any tumor progression (local, regional, or distant). The PFS rate will be estimated using the Kaplan-Meier (KM) method. The 95% CIs for median PFS and other quartiles will be estimated using the Brookmeyer-Crowley method. The PFS rates at 3, 6, 9, and 12 months will be estimated by the Kaplan-Meier method and their corresponding 95% CIs will be calculated using Greenwood's formula for estimating the variance and log-log transformation with back transformation to CIs on the untransformed scale. The censoring rule is described as below and frequency of subjects (n, %) experiencing PFS events and censored by reason will be provided.

For subjects who are alive and without tumor progression (locally, regionally, distantly) during the study follow-up period, the subject will be censored at the date of the last tumor assessment.

Listings will be provided.

8.10.6 Overall Survival

Analyses will be performed on the FAS and at subject level. On run-in set, only the number and percentage of subjects who have died will be provided.

Overall Survival (OS) is defined as the time that the subject is still alive after the original ablation procedure within the study duration. The OS rates will be estimated using the Kaplan-Meier method. The 95% CIs for median OS and other quartiles will be estimated using the Brookmeyer-Crowley method. The OS rates at 1, 3,6, 9 and 12 months will be estimated by Kaplan-Meier method and their corresponding 95% CI will be calculated using Greenwood's formula for estimating the variance and log-log transformation. The censoring rule is described as below and frequency of subjects (n, %) experiencing death events and censored by reason will be provided.

For each subject who is not known to have died during the study follow-up period, the subject will be censored at last contact date (e.g., last contact date of AE end date, lesion assessment date, visit date, study completion date, or last known alive date).

8.10.7 Safety Analyses

Analyses will be performed at subject level on the SAF. The AE summary will be provided on the FAS and run-in set as well.

All AEs that occur on or post the initial ablation date or occurs before the initial ablation date and related (Unlikely, Possibly, Probably, or Causal) to the study device or study procedure will be presented.

AEs will be coded using MedDRA. The number of events and the number and percentage of subjects reporting AEs and SAEs will be summarized at the MedDRA system organ class and preferred term level. The severity of AEs and SAEs will be summarized. The AEs and SAEs will also be summarized according to the Society of Interventional Radiology (SIR) classification. Similar summaries will also be provided for AEs and SAEs related to the study device, AEs and SAEs related to the study procedure, and device deficiencies. Related events are those where the relationship is indicated as Unlikely, Possibly, Probably, or Causal by the PI. In addition, related events where the relationship is indicated as Possibly, Probably, or Causal will also be summarized. Summaries of all AEs and SAEs reported within the first 30 days after the initial ablation procedure and overall will be generated. The incidence of predefined AEs will also be summarized and reported, including pneumothorax (all and those requiring chest tube drainage), hemorrhage (all and those requiring treatment), post-ablation syndrome, chest wall pain, pleural effusion (all the those requiring chest tube drainage), pneumonia, pulmonary or pleural abscess, other infections, and bronchopleural fistula.

All AE/device deficiencies will be listed.

All abnormal laboratory tests and positive blood or urine pregnancy test results will be listed.

8.11 Exploratory Endpoints and associated hypotheses

No hypothesis testing is performed for the exploratory endpoints. Exploratory endpoints will be analyzed on the FAS and run-in set.

8.11.1 Ablation-Procedure Related Pain

Descriptive summary statistics of actual values and change from pre-ablation by post-ablation analysis time point (i.e., visit 2B, visit 3-6) will be presented by ablation-procedure related pain measured by Numeric Pain Rating Scale. The change from baseline will be tested by paired-t test. The number of missing, mean, SE, and 95% CI based on student's t distribution will be provided.

The shifting table of baseline vs. post-baseline will be presented. 0: No pain; 1-3: Mild pain; 4-6: Moderate pain; 7-10: Severe pain. The number of subjects and percentage will be provided.

8.11.2 Quality of life

The quality of life assessment included the EORTC QLQ-C30 (version 3.0) and the lung-specific EORTC QLQ-LC13 (Figure1 and Figure2). The EORTC QLQ-C30 is scored for global health status/QOL and 5 functional (physical, role, emotional, cognitive, and social) and 3 symptom (fatigue, nausea and vomiting, and pain) scales. Additionally, 6 single-item scales are included (dyspnea, appetite loss, constipation, diarrhea, and financial difficulties). The EORTC QLQ-LC13 module incorporates 1 multi-item scale to assess dyspnea and a series of single-item scales assessing coughing, pain, sore mouth, dysphagia, peripheral neuropathy, alopecia, and hemoptysis. For both the EORTC QLQ-C30 and QLQ-LC13, raw scores are converted into scale scores ranging from 0 to 100. For the GHS/QOL and functional scales, higher scores represent better HRQOL; for the symptom scales, lower scores represent fewer symptoms. Descriptive statistics of all scale scores over time and change from pre-ablation will be provided for EORTC QLQ-C30 and QLQ-LC13, respectively.

The change from baseline to each post-ablation visit will be tested by paired-t test. The number of missing, mean, standard error (SE), and 95% CI based on student's t distribution will be provided.

Figure1 Scoring the QLQ-C30 version 3.0

	Scale	Number of items	Item range*	**Version 3.0** Item numbers	Function scales
Global health status / QoL					
Global health status/QoL (revised)†	QL2	2	6	29, 30	
Functional scales					
Physical functioning (revised)†	PF2	5	3	1 to 5	F
Role functioning (revised)†	RF2	2	3	6, 7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social functioning	SF	2	3	26, 27	F
Symptom scales / items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnoea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhoea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix "2" – for example, PF2.

Figure2 Scoring the QLQ-LC13

Scoring of the lung cancer module

The lung cancer module incorporates one multi-item scale to assess dyspnoea, and a series of single items assessing pain, coughing, sore mouth, dysphagia, peripheral neuropathy, alopecia, and haemoptysis.

The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales / single items of the QLQ-C30.

Scale name	Scale	Number of items	Item range*	QLQ-LC13 Item numbers	†
Symptom scales / items					
Dyspnoea†	LCDY	3†	3	3,4,5	X
Coughing	LCCO	1	3	1	
Haemoptysis	LCHA	1	3	2	
Sore mouth	LCSM	1	3	6	
Dysphagia	LCDS	1	3	7	
Peripheral neuropathy	LCPN	1	3	8	
Alopecia	LCHR	1	3	9	
Pain in chest	LCPC	1	3	10	
Pain in arm or shoulder	LCPA	1	3	11	
Pain in other parts	LCPO	1	3	12	

* "Item range" is the difference between the possible maximum and the minimum response to individual items.

† The dyspnoea scale should only be used if all three items have been answered. Some respondents ignore question 5 because they never climb stairs; in this case, the score for the dyspnoea scale would be biased if it were based upon the other two items. Hence if item 5 is missing then items 3 and 4 should be used as single-item measures.

8.11.3 ECOG performance status

Descriptive summary statistics of actual values and shift from pre-ablation by post-ablation analysis time point (i.e., visit 2B, visit 3-6) will be presented by ECOG performance status score. Number of subjects and percentage will be provided.

8.11.4 Ablation procedure related indicators

The following analyses will be performed on the tumor level:

Selected information from the ablation procedure (number of ablation cycles, duration of ablation for each cycle, total duration of ablation, maximum power applied, maximum temperature observed, and number and types of probes used) will be summarized for the initial ablation and re-ablations, respectively, with descriptive statistics.

The following analyses will be performed on the subject level：

Duration of procedure, mode of anesthesia, and guidance method used will be summarized for the initial ablation and re-ablations, respectively, with descriptive statistics.

Listings will be provided.

8.11.5 Length of hospital stay

Length of Hospital Stay is defined as the difference of date of discharge and date of completing the ablation procedure. Descriptive summary statistics of the length of stay will be presented for the initial ablation and re-ablations.

8.11.6 Re-admission rate

The Re-admission Rate is defined as the percentage of subjects with any re-admission to the hospital within 30 days post-ablation (including the original ablation or a repeat ablation). The number and percentage of subjects with re-admissions will be summarized, along with the reasons for re-admission.

Listings will be provided for those with any re-admission.

8.12 Plans for Interim Analysis

No interim analyses are planned for the purpose of stopping the study early for success or futility.

An interim analysis for study sample size re-estimation per the observed Technical Efficacy Rate will be performed after 50% of the target enrolled subjects complete the Visit 3 (at least n_1=54 evaluable subjects with primary effectiveness endpoint data). To warrant the originally intended 80% power of the study at the 0.025 level (1-sided), the sample size may be increased to a maximum of 200 enrolled subjects in order to provide 180 evaluable subjects, given the expected dropout rate.

The sample size re-estimation method is based on evaluation of conditional power in relationship to pre-specified decision criteria defined by ranges of attainable conditional power values. Only an increase in sample size is possible under this approach when observed conditional power falls within the 'promising zone' as described below (Bhatt 2016, Chen 2004, Cui 1999, Mehta and Pocock 2011, Orloff 2009).

The independent statistician will review the conditional power, and make a recommendation per pre-specified criteria about the sample size increasing or maintaining the originally planned sample size.

The sponsor is not allowed to access the interim results prior to the primary analysis.

8.12.1 Conditional Power

Conditional power refers to the probability of concluding a positive study at the end of trial, given interim results of treatment effect and the assumed true treatment effect. The conditional power at the interim analysis $CP_{\hat{\delta}_1}(z_1, \tilde{n}_2)$ can be defined by the following equation from Mehta and Pocock (2011):

$$CP_{\hat{\delta}_1}(z_1, \tilde{n}_2) = 1 - \Phi\left(\frac{z_\alpha\sqrt{n_2} - z_1\sqrt{n_1}}{\sqrt{\tilde{n}_2}} - \frac{z_1\sqrt{\tilde{n}_2}}{\sqrt{n_1}}\right)$$

Where $\tilde{n}_2 = n_2 - n_1$, $z_\alpha = \Phi^{-1}(1-\alpha)$, and z_1 is the value of test statistic as computed at the interim analysis and $Z_1 = \frac{P_1 - P_0}{\sqrt{P_0 \times (1-P_0)/n_1}}$. Define the observed interim estimate, $\hat{\delta}_1 = \hat{p}_1 - p_0$, where $\hat{p}_1$ is Technical Efficacy Rate observed at interim analysis and p_0 is the PG.

Our targeted conditional power at the first interim review is 0.80 (i.e., same as the study power), and we pre-specify a range of conditional power values below 0.80 that would deem our interim results promising and warrant a sample size re-estimation. The "promising zone" for conditional power was defined as between 0.387 and 0.80, as per Mehta and Pocock (2011) method using one-sided alpha of 0.025 and power of 0.80. We propose the following criteria for sample size increase, depending on the zone into which $CP_{\hat{\delta}_1}(z_1, \tilde{n}_2')$falls at the interim look.

- Favorable zone: If $CP \geq 0.8$ (or equivalently $z_1 \geq 1.806$), continue to $n_2 = 108$
- Promising zone: If $0.387 \leq CP < 0.8$ (or equivalently $1.243 \leq z_1 < 1.806$) increase sample size to $n_2^* = \min(n_2', n_{max})$, $where\ n_2'\ is\ such\ that\ CP = 0.8$
- Unfavorable zone: If $CP < 0.387$ (or equivalently $z_1 < 1.243$), continue to $n_2 = 108$

The following table provide the parameters used for the sample size re-estimation.

Design parameters	**Value**
Interim Sample Size, n_1	54
Cumulative Final Sample Size, n_2	108
Incremental Sample Size, $\tilde{n}_2$	54
Pre-specified Maximum Allowable Sample Size, n_{max}	180
Drop-out rate	10%
Significance Level (Alpha)	0.025 (one-sided)
Re-Estimated Cumulative Final Sample Size, n_2^*	Derived below
Re-Estimated Incremental Sample Size, $\tilde{n}_2'$	Derived below

The re-estimated cumulative final sample size (n_2^*) is computed as follows:

$n_2^* = \min(n_2', n_{max})$, where $\tilde{n}_2'$ satisfies the condition $CP_{\hat{\delta}_1}(z_1, \tilde{n}_2') = 1 - \beta, \beta = 20\%$, and n_2' is the estimated new sample size computed as $n_1 + \tilde{n}_2'$.

Based on Mehta and Pocock (2011), this condition is satisfied by the function:

$\tilde{n}_2' = \left[\frac{n_1}{z_1^2}\right]\left[\frac{z_\alpha\sqrt{n_2}-z_1\sqrt{n_1}}{\sqrt{n_2-n_1}}+z_\beta\right]^2$, where $z_\alpha = \Phi^{-1}(1-\alpha)$, $z_\beta = \Phi^{-1}(1-\beta)$, and z_1 is the test statistics as computed at the interim analysis of $n_1 = 54$ evaluable subjects.

Details of $\hat{p}_1$, z1, CP, Re-Estimated Incremental Sample Size, Re-Estimated Cumulative Final Sample Size

$\hat{p}_1$	z1 at interim	CP	$\tilde{n}_2'$	n_2^*
0.868	1.243	0.387	197	180

8.13 Plans for Primary and Final Analysis

Primary analysis will occur after all enrolled subjects finish the 1-month visit, aiming to evaluate Technical Efficacy Rate, and periprocedural endpoints during the 1st month after initial lung tumor ablation. If the primary effectiveness endpoint success criteria is met, the study is considered as a success and the study report will be submitted to NMPA for the registration application. Per the imputation rule for the sensitivity analysis of the primary endpoint, if the subject does not have a Visit 3, then chest enhanced CT at Visit 4 will be sent to the IRC for local tumor progression assessment, and the primary effectiveness endpoint at Visit 3 will be imputed per if local tumor progression occurs by Visit 4. Therefore, if there is any imputation that requires the evaluation from Visit 4, the primary analysis will occur after these subjects' visit-4 results are ready or confirmed as missing.

A final analysis will be performed after all subjects have completed the study or early discontinued from the study.

8.14 Handling of Missing Data

For missing data, the number of missing be displayed. All summaries will be performed only on subjects undergoing ablation with the NeuWave Certus Microwave Ablation System. A new subject will be recruited to replace any subject who discontinues the ablation procedure after the ablation probe punctures the skin and before the microwave energy is initiated. If an initial ablated tumor is re-ablated before the Visit 3 Technical Efficacy assessment, the re-ablated tumor will be considered as a failure.

- In the primary analysis of the primary efficacy endpoint: other subjects with missing primary efficacy data or other early discontinuations will not have data imputation.
- In the sensitivity analysis of the primary efficacy endpoint: as a conservative surrogate endpoint, if the subject does not complete Visit 3, then chest enhanced CT at Visit 4 will be sent to the IRC for local tumor progression assessment, and if local tumor progression occurs, the primary efficacy endpoint at Visit 3 will be imputed as a failure.

If there is no local tumor progression at Visit 4, the primary efficacy endpoint at Visit 3 will be imputed as a success. If a subject does not have both Visit 3 and Visit 4, the subject will be excluded from the analysis of the primary efficacy endpoint.

No imputations will be made for missing demographics, baseline characteristics, tumor characteristics, or secondary and exploratory endpoint variables in the analyses.

For time-to-event analysis, all censored data will be accounted for using appropriate statistical methods as described in sections 8.10.4-8.10.6.

8.15 Subgroup Analysis

Analyses will be done on the FAS for exploratory purpose appropriately.

Descriptive subgroup analyses of the following effectiveness endpoints at the tumor level, including but not limited to Technical Efficacy Rate, and technical success rate.

The number, percentage, and its two-sided 95% Clopper-Pearson CI will be provided.

The following subgroup variables will be considered:

- Tumor type: Stage IA1 NSCLC, Stage IA2 NSCLC, Oligometastatic lung tumor, Other
- Sites

8.16 Assessment of Site Homogeneity

Analyses will be done on the FAS

The technical efficacy rate and technical success rate endpoints will be descriptively summarized by site.

9 Data Monitoring Committee (DMC)

The use of a DMC is not planned. However, if a DMC is appointed, details will be documented within a DMC charter.

Reference:

- Bhatt DL, Mehta C. Adaptive Designs for Clinical Trials. N Engl J Med. 2016 Jul 7;375(1):65-74. doi: 10.1056/NEJMra1510061. PMID: 27406349
- Chen YH, DeMets DL, Lan KK. Increasing the sample size when the unblinded interim result is promising. Stat Med. 2004 Apr 15;23(7):1023-38. doi: 10.1002/sim.1688. PMID: 15057876.
- Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. Stat Med. 2011 Dec 10;30(28):3267-84. doi: 10.1002/sim.4102. Epub 2010 Nov 30. PMID: 22105690.
- Orloff J, Douglas F, Pinheiro J, et al. The future of drug development: advancing clinical trial design. Nat Rev Drug Discov. 2009 Dec;8(12):949-57. doi: 10.1038/nrd3025. Epub 2009 Oct 9. PMID: 19816458.

End of Document